DOCUMENT RESUME

ED 282 230 CS 210 519

AUTHOR Chafe, Wallace; Danielwicz, Jane
TITLE Properties of Spoken and Written Language. Technical Report No. 5.
INSTITUTION California Univ., Berkeley. Center for the Study of Writing.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO CSW-TR-5
PUB DATE May 87
NOTE 31p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS English; Higher Education; Language Research; *Language Usage; *Oral Language; *Reading Writing Relationship; Sentence Structure; Vocabulary Skills; *Writing Processes; Writing Research; *Written Language

ABSTRACT

To find differences and similarities between spoken and written English, analyses were made of four specific kinds of language. Twenty adults, either graduate students or university professors, provided a sample of each of the following: conversations, lectures, informal letters, and academic papers. Conversations and lecture samples came from tape recordings of casual dinner table talk and from class lectures. Letters and papers were voluntarily provided by the subjects. Findings indicated that conversationalists employed a relatively limited vocabulary, made inexplicit lexical choices, and used many colloquialisms. They created relatively brief intonation units which they chained together, stopping occasionally to form a sentence boundary. Conversationalists also interacted with their audiences, showed ego involvement, and often spoke specifically. Academic lecturers employed an equally limited vocabulary, also used hedges and were referentially inexplicit. They used literary vocabulary more than conversationalists, but their intonation units were only slightly longer and they created no more syntactically complex sentences. In contrast, letter writers used more varied vocabulary, and only sometimes hedged or were inexplicit. They used a moderate number of colloquial words and contractions, but many more literary items. Their intonation units were intermediate between conversationalists and academic writers, and their sentences were usually well formed. Letter writers were more involved with themselves and with concrete reality than any other language users in the study. Academic writers exhibited the extremes of writing behavior. They used large varied vocabulary and hedges, inexplicit references, colloquialisms, and contractions. Academic writers used maximally long intonation units, expanded them with all available syntactic devices, and were detached from both concrete reality and themselves. (SKC)

ED282230

Center for the Study of Writing

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

University of California, Berkeley
Carnegie Mellon University

CS 210 519

CENTER FOR THE STUDY OF WRITING

Technical Report No. 5

PROPERTIES OF SPOKEN AND WRITTEN LANGUAGE

Wallace Chafe
Jane Danielwicz

May, 1987

To appear in Comprehending Oral and Written Language edited by R. Horowitz and F.J. Samuels. New York: Academic Press. 1987

University of California
Berkeley, CA 94720

Carnegie Mellon University
Pittsburgh, PA 15213

The project presented, or reported herein, was performed pursuant to a grant from the Office of Educational Research and Improvement/Department of Education (OERI/ED) for the Center for the Study of Writing. However, the opinions expressed herein do not necessarily reflect the position or policy of the OERI/ED and no official endorsement by the OERI/ED should be inferred.

CENTER FOR THE STUDY OF WRITING

Director	Sarah Warshauer Freedman University of California, Berkeley	
Co-Directors	Linda Flower Carnegie Mellon University	J.R. Hayes Carnegie Mellon University
	James Gray University of California, Berkeley	
Assistant Director	Sandra Murphy University of California, Berkeley	

Publication Review Board

Chair	Melanie Sperling University of California, Berkeley	
Assistant Chairs	Charles Elster University of California, Berkeley	Karen Schriver Carnegie Mellon University
Advisor	Jill H. Larkin Carnegie Mellon University	

Carla Asher, *Herbert H. Lehman College of the City University of New York*
Nancie Atwell, *Boothbay Region Elementary School, Boothbay Harbor, ME*
Robert de Beaugrande, *University of Florida*
Ruby Bernstein, *Northgate High School, Walnut Creek, CA*
Wayne Booth, *University of Chicago*
Robert Calfee, *Stanford University*
Michael Cole, *University of California, San Diego*
Colette Daiute, *Harvard University*
John Daly, *University of Texas, Austin*
Peter Elbow, *State University of New York, Stony Brook*
JoAnne T. Eresh, *Writing and Speaking Center, Pittsburgh, PA*
Donald Graves, *University of New Hampshire*
James Hahn, *Fairfield High School, Fairfield , CA*
Julie Jensen, *University of Texas, Austin*
Andrea Lunsford, *Ohio State University*
Marion M. Mohr, *Fairfax County Public Schools, Fairfax County, VA*
Lee Odell, *Rensselaer Polytechnic Institute*
Charles Read, *University of Wisconsin*
Victor Rentel, *Ohio State University*
Michael W. Stubbs, *University of London*
Deborah Tannen, *Georgetown University*
Gordon Wells, *Ontario Institute for Studies in Education*

PROPERTIES OF SPOKEN AND WRITTEN LANGUAGE

By

Wallace Chafe and Jane Danielwicz
University of California, Berkeley

There can be no doubt that people write differently from the way they speak. Current interest in the nature of the differences between "orality" and "literacy" was stimulated two decades ago by Jack Goody and Ian Watt (1968), who combined the insights of an anthropologist and a literary scholar to give us a perspective on the consequences of literacy. In linguistics, Josef Vachek (1973) brought together the results of his and the Prague School's pioneering concern for the special nature and importance of written language. In educational psychology, David Olson (1977) explored the consequences of the "essayist tradition" in writing. These classic writings, and a number of other works which have lately been appearing with increasing frequency, have signaled the end to a period in which the systematic study of language was dominated by Leonard Bloomfield's dictum to the effect that "writing is not language, but merely a way of recording language by means of visible marks" (Bloomfield, 1933:21).

The range of viewpoints and findings has continued to expand. A brief but representative cross-section of recent works might include the volumes written or edited by Kroll and Vann (1981), Scribner and Cole (1981), Tannen (1982a), and Ong (1982). Volumes like these show a concern for a range of topics which includes the kinds of language that are produced in speaking and writing, the cognitive and social reasons why spoken and written language may differ, the uses people make of speaking and writing, and the different effects spoken and written language may have on the way people think. Our own earlier work (Chafe, 1982, 1985, 1986) tried to identify more precisely the differences to be found in the kinds of language which are used by speakers and writers, and we speculated on underlying causes for those differences. These earlier reports of ours were restricted to a comparison of what we supposed were two extremes of "spokenness" and "writtenness": conversational speaking at the one extreme, and academic writing at the other.

It has always been clear, however, that neither spoken language nor written language is a unified phenomenon. Far from there being one single kind of language that people speak and one other kind that they write, each of these two modes itself allows a multiplicity of styles. But, beyond that, there is a great deal of overlap between speaking and writing, in the sense that some kinds of spoken language may be very written-like, and some kinds of written language very spoken-like (Tannen, 1982b). In examining varied samples of spoken and written language, one may even be led to wonder whether that distinction makes any sense at all, or whether there are just many varieties of a language which are available to its speakers, most or all of which varieties may be either spoken or written depending on the circumstances.

Certainly there are a number of factors responsible for differences in the kinds of language a person may use. And certainly one of these factors is the matter of whether the language is produced with the mouth and received with the ear, or whether it is produced with the hand and received with the eye. We will be looking here at some of the differences in language which seem to have much to do with that difference in how it is produced and received. At the same time we will be looking at how some of the uses to which language is put interact with the spoken-written distinction. The context of language use, the purpose of the speaker or writer, the subject matter of what is being said or written--these are some of the other factors which influence the form language takes. It is instructive to notice that the many languages which have never been written at all also exhibit different styles--sometimes radically different styles--which are used for different purposes (see in this connection Chafe, 1981). Writing increases the ways in which language can be used, and adds significantly to the linguistic repertoire. But of course aspects of written style may be borrowed by speakers when it suits them, just as aspects of spoken style may be borrowed by writers. In what follows we will be taking particular note of the varying proportions of linguistic features which surface as both spoken and written language are used in different circumstances.

Our observations come from a project in which we collected four kinds of language from each of 20 adults.[1] All of them were either professors or graduate students at the University of California at Berkeley or the State University of New York at Albany. Our reason for investigating styles of language among this very specialized population was that we wanted to compare four specific kinds of language that are produced in the course of a person's normal activities. Academic people happen to produce these four kinds quite naturally. We will refer to the four as conversations, lectures, letters, and academic papers.

The conversational samples were taken from tape recordings of casual talk which took place during, immediately preceding, or immediately following an informal dinner in one of our homes. Each subject was invited separately to a dinner, with the understanding that the conversation was to be tape-recorded. From each of the resulting several hours of taped conversation, we selected a portion consisting of at least 100 "intonation units" produced by that subject. An intonation unit can be thought of for the moment as a single clause, but its nature will be discussed more fully below. We selected a portion of the tape during which the speaker in whom we were interested did most of the talking. As a result, most of the samples consist of one or more narratives of personal experiences. We transcribed these portions, recording not only the words uttered but also disfluencies, laughter, and, in a minimal way, intonation. We then proceeded to code these samples for the features which we will discuss presently.

The lecture samples were taken from tape recordings of these same professors or graduate students speaking to classes--in all cases small classes addressed under relatively informal circumstances. That is, we did not record formal lectures presented to large audiences. Again we selected at least 100 intonation units from each recorded sample, trying to choose a portion that was representative of the subject's lecturing style. We transcribed and coded these samples in the same manner as the others.

The samples of letters consisted of informal letters written by the same professors or graduate students to relatives, friends, or colleagues. In the end we found these materials to be somewhat less satisfying than the other three types, for several reasons. They were harder to obtain in the first place, since the subjects differed in their

eagerness and ability to provide us with letters, and those who habitually do most of their informal communicating by telephone simply found it hard to come up with any. The samples also tended to be shorter than those of the other types, some of them consisting of fewer than 100 idea units. They also differed considerably in tone and purpose, and were in general less homogeneous than the other kinds of language we collected. In spite of these reservations, we believe that the letters represent a reasonable sample of the informal letter-writing style of academic people. We recopied each letter, dividing it into intonation units according to punctuation (see below) and coded it as we did the other materials.

Finally, the academic writing samples were excerpts from articles that had been written for academic journals or books. They were usually volunteered by their authors as representative of their academic writing styles. We selected portions of at least 100 intonation units each, and processed them in the same manner as the other three kinds of language.

We found, not surprisingly, that these four styles of language--two of them spoken and two of them written--were quite different from each other. Some of the differences appeared to be caused only by the fact that the language was spoken or written. More often, there were additional factors of language use which interacted with the spoken-written distinction.

We will begin by focusing on how speakers and writers choose words and phrases appropriate to what they want to say. The relevant point here is that speakers must make such choices very quickly whereas writers have time to deliberate, and even to revise their choices when they are not satisfied. As a result, written language, no matter what its purpose or subject matter, tends to have a more varied vocabulary than spoken. Another aspect of lexical choice is the style of the vocabulary items which are selected, the degree to which words and phrases are colloquial or literary. Here it is easily possible for lecturers to borrow from literary vocabulary, and for letter writers to colloquialize their product by drawing on the vocabulary of spoken language. The division between spoken and written language is therefore not as sharp in this respect, since the operative constraints are not associated with an inevitable factor such as speed of processing, but involve stylistic decisions which are easily transferrable from one mode of language production to the other.

We turn next to the matter of how words are put together in clauses. We note that the relevant unit of spoken language appears to be the basically prosodic entity we have been calling an intonation unit, and this notion is explained. Intonation units are found to be longer in written than in spoken language, a fact that can again be attributed to different processing constraints. Spoken intonation units are limited in size by the short-term memory or "focal consciousness" capacity of the speaker, and perhaps also by the speaker's awareness of the listener's capacity limitations. Writers are under no such constraints, and as a consequence they have a tendency to produce expanded intonation units. We go on to consider a few of the many devices writers use to create this expansion.

Next we pay some attention to the different ways intonation units are joined together to form sentences in spoken and written language. Spoken language relies to a large extent on a chaining technique, and avoids elaborate syntactic relations among clauses. We point out that intonation units appear to be the natural unit of speaking, whereas integrated, elaborated sentences have become the natural unit of writing.

We turn finally to features of language related to the social interaction which is natural to speaking, as contrasted with the social isolation which is inherent in writing. We distinguish between involvement with the audience, involvement with oneself, and involvement with concrete reality. We go on to show how detachment from these kinds of involvement is manifested in academic writing. We speculate that writing makes possible a kind of abstract thinking that is less than normal in the conversational use of language, but that is not a necessary component of writing either.

We will proceed now with a closer examination of these points. As we examine each in turn, it will be helpful to keep in mind the distinction between those differences which exist because of differences in the speaking and writing processes themselves, and those differences which have arisen because of the varied contexts, purposes, and subject matters of both spoken and written language. We will show how differences of this latter kind are more easily overridden when the uses of speaking and writing overlap.

VARIETY OF VOCABULARY

One thing that speakers and writers alike must do is to choose words and phrases that will express what they have in mind. They must find the vocabulary to convey their thoughts. We assume there is a discontinuity between what people have in mind and the language they use to express it. What is going on in people's heads does not always translate automatically into appropriate words and phrases. Knowledge of a language includes knowledge of a huge repertoire of lexical options that can be used for referring to the objects, states, and events which people may have occasion to talk about. As people construct language, they constantly have to choose among these options in ways they hope will communicate appropriately what they are thinking of. It would thus appear that the choice of effective ways of saying things must call for the expenditure of some cognitive effort.

Of interest here is the fact that writing provides more time, and ultimately more resources, for this effort. The way we choose words and phrases in speaking and the way we choose them in writing are not the same processes at all. Speaking is done on the fly, while writing is both slow and editable. When we speak we have little time to choose our words, and once we have uttered them, they have been uttered. If we are not satisfied, we may try to revise what we have said, but too much fumbling is harmful to effective communication, and in any case our fumbling is laid bare for all to hear. When we write, we can produce language at whatever pace we wish. We can take hours, if we need to, to find an appropriate word. And there is no need to remain committed to the first lexical choices we make. Whatever words and phrases we may initially decide on, we are free to revise them again and again until they satisfy us. We have the leisure to dip into the rich storehouse of literary vocabulary, searching for items that will capture nuances which, if we were speaking, we would not have time to bother with. And this whole editing process is hidden from the eventual consumer of our language, to whom we can pretend that the aptness of our choices flowed naturally from our pen, typewriter, or word-processor.

As a consequence of these differences, speakers tend to operate with a narrower range of lexical choices than writers. Producing language on the fly, they hardly have time to sift through all the possible choices they might make, and may typically settle on the first words that occur to them. The result is that the vocabulary of spoken language is more limited in variety. A mechanical measure of this difference is the

type/token ratio of words: the number of different words in a sample divided by the total number of words in that sample. As a simple illustration, the first sentence in this paragraph contains 19 word tokens but only 17 word types, since the words "a" and "of" both occur twice. The type/token ratio of this one sentence is thus 17/19, or .89. It should be noted that such a ratio decreases as the number of words in a sample increases, so that the ratio just given cannot be compared with those given for our complete samples in Table 1. It can be seen from Table 1 that written language consistently shows a higher ratio than spoken, and that both styles of speaking (conversations and lectures) have lower type/token ratios, and thus less lexical variety, than both styles of writing (letters and academic papers).

Table 1. Type/Token Ratios

Conversations:	.18
Lectures:	.19
Letters:	.22
Academic Papers:	.24

It is of some interest that the ratio in lectures is about the same as that in conversations. It appears that the necessarily rapid production of spoken language consistently produces a less varied vocabulary, regardless of the kind of speaking involved. Thus there seems to be a limit on the lexical variety a speaker can produce under any circumstances, a limit which cannot be overridden by a change in context or purpose. In speaking one always has to choose words quickly. In writing, with or without editing, one always has more time for choice. In other words, the constraints inherent in the speaking and writing processes are dominant here, and are not overridden by different uses of the two.

There is evidence that speakers are aware of their limitations in choosing lexical items which adequately convey what they have in mind. An indication of such awareness is the use of hedges like "sort of" or "kind of," as in the following examples:

(1) ... And ... she was still young enough so I .. I just .. was able to put her in an .. uh--*sort of* ... sling.. I mean one of those tummy packs .. you know, (conversation)

(2) A--nd the graduate students are *kind of* scattered around, (conversation)

(3) ... Um ... I'm *sort of* I'm paraphrasing a little bit (lecture)

(4) ... So ... this one is *kind of typical*. (lecture)

Such expressions suggest that, although these speakers settled quickly on the lexical choices "sling," "scattered around," "paraphrasing," and "typical," they were not completely satisfied with these ways of expressing things (cf. Lakoff, 1975). The speaker of (1), for example, indicated with "sort of" that what she was thinking of was not a typical sling, as she subsequently confirmed with her switch to "tummy pack" as a possibly

better choice. Table 2 shows the distribution of hedges among our four kinds of language. The figures show the mean number of occurrences per thousand words in each sample. These figures are not large for any of the four styles, but there is a clear difference between the two spoken styles, where the incidence of hedges was identical, and the two written styles, where hedges were hardly present at all.

Table 2. Hedges

Conversations:	4
Lectures:	4
Letters:	1
Academic Papers:	0

Type-token ratios and the use of hedges both have to do with the adequacy with which words express people's underlying intentions. Another, similar feature is the degree of referential explicitness. Speakers not only have less time to choose vocabulary, but they also cannot or do not take the time to be as explicit about what they are referring to. A symptom of this kind of vagueness is the use of third person neuter pronouns, usually "it," "this," or "that." Typically, the antecedent of a pronoun has been spelled out in an earlier noun phrase. Sometimes, however, and especially in speaking, there is no such clear antecedent:

(5) And they have been arguing (laugh) about this
for ... I don't know!
.. At least fifteen years or so.
... And ... *it* seems like a fairly fruitless
debate. (lecture)

The pronoun "it" in the last line refers to what was talked about in the two preceding lines (the fifteen-year-long argument), but that antecedent was never spelled out in an explicit noun phrase. Similarly, the pronoun "this" in the last line of (6) refers to the entire situation described in the lines above it:

(6) ... The fact that they're not ... alike,
... shows that ... language .. m ... can ..
forces an assignment ... of one of those objects
.. as reference point,
.. with respect to the other object,
... so the second object .. acts as reference point,
... I'm calling it the ground object,
... a--nd the first one is relatively figure to it,
... its ... path .. or site ... is characterized with
respect to it.
... *This* happens in time as well, (lecture)

Likewise, the pronoun "that" in the last line of (7) refers to the idea that husbands and wives would understand each other and have much in common, again an idea scattered over the preceding lines:

(7) ... And I picked husbands and wives because I thought
... first of all
... we would think that husbands and wives would
understand each other.
... if anybody can.
... Because they live together,
and they .. have so much in common.
... Um although *that* often turns out not to be the
case. (lecture)

In Table 3 we see the distribution of such uses of pronouns in our four types of language, again listed in terms of the mean number of occurrences per thousand words. It can be seen that this phenomenon too is almost equally present in both styles of spoken language, whereas its frequency is considerably lower in either style of writing. In academic papers there are very few examples of it.

Table 3. Inexplicit Third Person Reference

Conversations:	24
Lectures:	22
Letters:	11
Academic Papers:	4

We have suggested, then, that the degree of richness of vocabulary, as measured by the type-token ratio, is limited by the rapidity of production of spoken language, whereas the added time and the editing possibilities allowed by writing increase the variety of lexical choice. And we have noted that there exists here a relatively pure difference between speaking and writing, one little influenced by the particular kinds of speaking or writing involved. There is here, in other words, a processing constraint which cannot easily be overridden by other factors. The same is true of certain related features, including the use of both hedges and inexplicit third person neuter pronouns. With regard to all these features relevant to adequacy of vocabulary, there is a sense in which written language might be thought to be a superior vehicle, just because it provides more time for a subtle and explicit packaging of thoughts. On the other hand, speaking has the advantage of providing a more direct expression of ongoing thought processes. For listeners, as well as for us as investigators, speaking makes available a more direct window on the mind in action.

LEVEL OF VOCABULARY

So far we have dwelled on the fact that speakers have less time than writers to choose lexical items adequate to what they want to convey. We need now to pay attention to the fact that speakers and writers do not choose from the same supply, that the store of words and phrases used in conversations is not the same as that used in formal writing, for example, and in general that spoken and written vocabularies are partially different. Many, perhaps all languages have several levels or registers which are used for different purposes and in different situations. Japanese, for example, is well known for dictating somewhat different lexical choices depending on the relative social status

of the interlocutors. Many languages have ritual vocabularies which are different from those used in ordinary conversation. (How many times do we use the English word "merciful" outside of ritual contexts?) It is not surprising that languages with a long written tradition should have developed partially different vocabularies for speaking and writing. Not only does spoken language use a smaller repertoire, but the spoken and written repertoires contain somewhat different items.

There may be various ways in which such differences have arisen, but probably a common way is through differing rates of lexical change, spoken language being in general more innovative and written language more conservative. Thus, the spoken vocabulary of a language is likely to contain both new words and new senses of old words that are not present in the written vocabulary. At the same time the written vocabulary is likely to have retained older words and older senses of words which have passed out of spoken use. New uses, originating in the spoken language, may eventually enter the written vocabulary and become perfectly at home there. Samuel Johnson thought that the word "civilization" was not properly literary, but we would hardly make such a judgment today.

Perhaps spoken language compensates for its restricted lexical variety by assigning a premium to freshness. For example, the colloquial way of saying that something is good or desirable has changed from decade to decade--or even from year to year--through a series of words like "swell," "cool," "neat," "far out," "awesome," and so on. Speakers seem unwilling to let such a concept be expressed by a word that is stale, and thus engage in a never ending process of vocabulary replacement. Freshness of vocabulary is less a value in writing, where the important thing is to have a fixed stock of many subtly different items to choose from. Spoken language achieves richness through constant change within a limited range of choices; written language achieves it through broadening that range.

For our purposes, if we disregard slang, profanity, and professional jargon, lexical items in our samples of English fall into three classes: colloquial, literary, and neutral. There are certain items like "kid," "bike," "figure out," and "bunch of" which occur quite naturally in casual spoken language, but which are generally out of place in writing unless the writer is trying to imitate speaking. There are other items like "ascertain," "optimal," "despite," and "constitute" which are typically used in writing, but which sound out of place in speaking. Most vocabulary items, on the other hand--like "house," "run," "show," and "try"--are neutral to this distinction, being equally at home in both spoken and written language. English speakers appear to be well aware of these differences, and to be able to judge with considerable agreement the class to which a particular item belongs. Three of us have made such judgments with respect to the data under discussion here. Table 4 shows the number of occurrences per thousand words of distinctly literary or distinctly colloquial vocabulary.

It should be noticed that the distribution here is different from that which we saw in Tables 1-3. There the major, if not the only difference was between speaking and writing. The purpose of the speaking or writing was only minimally, if at all, significant. Now we are looking at a distribution where there is a continuum from conversations at one extreme to academic papers at the other. Lectures and letters fall somewhere between, and there is less difference between these spoken and written varieties than there is within the two kinds of speaking, or within the two kinds of writing. It may be that differences in level of vocabulary had their genesis in the different uses of speaking and writing, but they are obviously not constrained by the speaking and writing

processes themselves. There is nothing in the nature of speaking which prevents a speaker from using literary vocabulary, and nothing in the nature of writing which prevents a writer from using colloquial vocabulary. The level of word choice can be

Table 4. Literary and Colloquial Vocabulary

	Literary Vocabulary	Colloquial Vocabulary
Conversations:	8	27
Lectures:	19	18
Letters:	25	16
Academic Papers:	46	1

varied in whatever ways speakers and writers find appropriate to their contexts, purposes, and subject matters. Hence the fact that colloquial and literary vocabulary are equally present in lectures. Hence also the fact that letters, but not academic papers, have a fairly large colloquial component.

There is another feature whose use is much like the use of colloquial vocabulary. Spoken language commonly employs contractions like "it's," "I'm," and "don't." Such items are rare in academic written language, as shown in Table 5, where their distribution can be seen to be similar to that shown for colloquial vocabulary in Table 4, except that there is a greater difference between lectures and letters.

Table 5. Contractions

Conversations:	37
Lectures:	29
Letters:	18
Academic Papers:	0

Thus, evidently, the constraints on speaking vs. writing play a somewhat stronger role in the use of contractions, with both kinds of spoken language showing significantly more of them than either kind of written language. If we put that difference aside, contractions can be regarded as further examples of innovative spoken vocabulary, innovations which the most formal kind of written language avoids altogether, but which more casual writing is more willing to accept.

To summarize what we have found so far, choosing lexical items is partly a matter of choosing aptly and explicitly, partly a matter of choosing the appropriate level. In the first case, the deliberateness and editability inherent in writing lead to a more richly varied, less hedged, and more explicit use of words. Speakers are so strongly constrained by their need to produce language rapidly, and by their inability to edit, that they are unable to imitate the lexical richness and explicitness of writing even when, as in lecturing, such qualities would be especially valued. In the second case, although the separate histories of spoken and written language have led to partially divergent vocabularies, it is not as hard for speakers to borrow liberally from the

written lexicon, or conversely for writers to borrow from the spoken. Thus lectures are more literary than conversations, and letters more conversational than academic papers. The constraints are not imposed by cognitive limitations, but by judgments of appropriateness.

CLAUSE CONSTRUCTION

Language consists of more than words, and there is more involved in speaking and writing than deciding which words to use. The obvious next question is how speakers and writers combine words and phrases into clauses. Do they form clauses in partially different ways, and, if so, are such differences imposed inevitably by the speaking and writing processes? Or are speakers and writers easily able to cross the boundary between spoken and written language when they find it appropriate, as we saw them able to do in choosing between colloquial and literary levels of vocabulary?

Instead of basing our discussion on clauses purely and simply, we have found it more realistic to proceed in terms of a slightly different kind of unit. One of the most noticeable and consistent properties of casual spoken language is that it is produced in relatively brief spurts. Although we have sometimes called these spurts "idea units" (see especially Chafe, 1980 for further discussion), here we will use the term "intonation unit." The prototypical intonation unit has the following properties: (1) it is spoken with a single, coherent intonation contour, (2) it is followed by a pause, and (3) it is likely to be a single clause. Although the majority of intonation units are clauses, some contain an entire verb-complement construction while others are no more than a prepositional phrase, or even just a noun phrase, or a syntactic fragment of some other kind.

We have speculated on the cognitive basis of intonation units. It is fruitful to consider them the linguistic expression of that particular knowledge on which a speaker is focusing his attention at the moment. We can say that an intonation unit expresses what is in the speaker's short-term memory, or "focus of consciousness," at the time it is produced. Intonation units thus provide evidence as to the nature and capacity of focal consciousness. It seems that, under normal conditions, a speaker does not, or cannot, focus attention on more information than can be expressed in about 6 words. It seems, furthermore, that the syntax of intonation units must be kept fairly simple, and that attempts to achieve any very high degree of syntactic complexity are likely to cause trouble: hesitations, false starts, repetitions, and other verbal disfluencies (cf. Pawley and Syder, 1976).

The following is an example of conversational language presented with each intonation unit on a separate line. Sequences of two or three dots indicate pauses:

(B) ... I just this year have ... dropped down to
teaching half time.
.. which is what I've always wanted.
... You know I'm happy about it.
... It's a ... terribly long commute,
... a--nd now I'm just going two days a week.
... And just teaching one course a quarter.
... Cause the regular ... teaching load for us is six
courses a year.

We assume that written language has a covert prosody which is analogous to that of spoken language: that both writers and readers assign pitch, stress, and pauses to language as they write and read it. We assume also that an important aspect of this prosody is the assignment of intonation unit boundaries to written language. We assume, finally, that these boundaries are at least reasonably well indicated by punctuation. We are currently studying the extent to which the "punctuation units" of writing can be identified with the intonation units of speaking, but for the purposes of the present discussion we will simply equate the two. Thus, when we use the term "intonation unit" with reference to writing we will be referring to stretches of language between punctuation marks.

Table 6 shows the increase in intonation unit size, measured in terms of words per intonation unit, which takes place as we move from conversations to lectures, then to letters, and finally to academic papers.

Table 6. Words per Intonation Unit

Conversations:	6.2
Lectures:	7.3
Letters:	8.4
Academic Papers:	9.3

The difference between the two spoken language types on the one hand and the two written types on the other is what we might expect. Writing frees writers from the constraint which keeps down the size of spoken intonation units. They need not limit the production of language to what can be focused on at one time, but can spend an indefinite amount of time constructing intonation units of any size. Writing frees intonation units from the limitations of short-term memory.

It is interesting to speculate on why a writer should keep intonation units within the bounds of about 9 words. Two answers have occurred to us. One is that written intonation units are a carryover from spoken language. Since spoken language has been with the human race for so long, and since it is still the kind of language we use most of the time, it is not surprising that it should influence the form of written language. But another factor may be a writer's sensitivity to the task of a reader. Suppose that a reader reads written language in more or less the same way that a hearer hears spoken language--that intonation units are a significant unit not only of production, but also of comprehension. If that is true, then writers most concerned with the readability of their product may provide their readers with moderately sized intonation units because they know intuitively that language so packaged will be easier to process. The four intonation units of (9) approximate those of spoken language. The principal intonation unit of (10) is much longer, and possibly less readable for that reason:

(9) There are exceptions to this general rule,
and they are related to festivals of the winter
solstice and death. Christmas ritual focuses upon
the Virgin Mary, who is synonymous with the moon
in their belief.

(10) Thus, subjects can better recognize the shape of an average chair and give more consistent and detailed accounts of what a chair is and how one interacts with a chair than they can when asked to perform similar tasks with respect to the superordinate category furniture.

Lectures show a mean intonation unit length of 7.3 words, only slightly higher than conversations. It would seem that the cognitive constraint which limits the size of spoken intonation units is difficult to overcome. Although there are some properties of academic writing which academic lecturers successfully borrow--literary vocabulary as well as other features to be discussed below--lecturers are unable to transcend very far the limitations on intonation unit formation imposed by the capacity of focal consciousness.

Letters show a mean intonation unit length of 8.4 words, midway between lectures and academic papers. They represent, then, a mixture of the limits characteristic of spoken language and the expansiveness allowed by writing. We might think of letter writers as language producers who take some advantage of the possibilities allowed by writing, but who are not interested in carrying these possibilities as far as academic writers. One reason for their restraint may be that letter writers typically write faster and with less editing than academic writers, and thus do not take the time necessary to expand intonation units to the same degree. Another reason may be that they try to maintain the casualness of spoken language, realizing that shorter intonation units make for easier reading.

In summary, we have identified clause-like units of spoken language which we have called intonation units. We have found that these units are relatively brief in casual spoken language, and have attributed their brevity to the fact that speakers can focus their consciousness on only a limited amount of material at one time. Since writers do not have to produce language on the fly as speakers do, they are freed from this constraint. As a consequence, written language which has undergone much planning and editing, as represented here by academic writing, shows intonation units that are markedly longer. Academic speaking, on the other hand, characteristically produces intonation units which are only slightly longer than those of casual speaking, a fact which suggests that the cognitive limitation on the size of spoken intonation units is difficult to overcome. Letter writers produce intonation units of intermediate length, avoiding the extreme of academic writing and aiming at something closer to spoken language.

There are many linguistic devices whose effect is to increase the size of intonation units, and which are used more by writers than by speakers. Here we will discuss only a few of the more common of them, considering how they are distributed among our four kinds of language, and attempting to draw some conclusions from those distributions. There are three such devices which are used much more frequently than any others: prepositional phrases, nominalizations, and attributive adjectives. All three occur with considerable frequency in all four of the kinds of language we collected, a fact which suggests that there are no strong cognitive constraints limiting their use. On the other hand, they are unusually frequent in academic writing, where they are by far the favorite ways for academic writers to expand the size of intonation units. They occur with approximately equal frequency in lectures and in letters.

The following examples show how academic writing often piles prepositional phrase upon prepositional phrase:

(11) Language change has occurred when the utterances *of* some members *of* that community have characteristics demonstrably different *from* those *in* utterances *of* previous generations.

(12) In order to account for the use *of* linguistic features *by* certain speakers *in* prolonged interaction, I taped two and a half hours *of* naturally occurring conversation *among* six participants *at* a dinner *in* 1978.

The distribution of prepositional phrases in our data is shown in Table 7, in terms of the number of occurrences per thousand words.

Table 7. Prepositional Phrases

Conversations:	53
Lectures:	88
Letters:	91
Academic Papers:	117

The fact that they are used so commonly, even in conversations, suggests that they are easy to use--that they exact no very high price in terms of cognitive capacity. Conversationalists, however, are much less apt to construct prepositional phrase sequences like those seen in (11) and (12), and are more apt to rely on prepositions which form close constructions with verbs, or even to construct idea units which consist of nothing but prepositional phrases. The following sequence exemplifies both of these tendencies:

(13) so he took us *to* some lake,
... *in* Quebec,

The first idea unit is built around the memorized sequence "take (someone) to (somewhere)." The second shows the speaker giving full attention to a prepositional phrase to the exclusion of anything else.

These observations suggest that it would be rewarding, not just to count occurrences of prepositional phrases, but to examine more closely how they are used in different kinds of language. We have not yet undertaken such a study, except to examine occurrences of prepositional phrase sequences--two or more such phrases juxtaposed within the same intonation unit. The occurrences of such sequences per thousand words are shown in Table 8. Their predominance in academic writing is obvious, and again it is of interest that they are equally frequent in lectures and in letters.

Table 8. Prepositional Phrase Sequences

Conversations:	6
Lectures:	14
Letters:	14
Academic Papers:	22

Assuming that prepositional phrases are relatively easy to use, why should they occur with equal frequency in lectures and letters? We suggest that the lack of cognitive strain must make it possible for lecturers to use them more or less effortlessly, thereby mimicking written academic style. Still, the fact that lecturers are speaking and not writing prevents them from using prepositional phrases as abundantly as academic writers. Conversely, letter writers, writing more quickly and mimicking conversational style, do not make as much use of prepositional phrases as academic writers, but the extra time available to writers of any kind allows even letter writers to use them more than conversationalists. Thus the opposite goals and constraints of lecturers and letter writers lead to an almost identical result.

A second favorite device used by writers for increasing the length of intonation units is nominalization, the formation of a noun from a verb, as with the words "representation," "determinant," and "performance" in the following examples:

(14) It is at this level of language that the *representation* of space is explored here.

(15) If processing time is a crucial *determinant* of memory *performance*,

Nominalizations are a principal means whereby a single clause can be constructed from what might otherwise have been several clauses. For example, in place of (14) a conversationalist might have said something about "how people represent space," increasing the number of intonation units by adding one based on the verb "represent." The distribution of nominalizations is similar to that of prepositional phrases, though they are slightly less frequent, as shown in Table 9.

Table 9: Nominalizations

Conversations:	27
Lectures:	56
Letters:	55
Academic Papers:	92

Many words which originated as nominalizations have become standard items of the academic vocabulary, without which academic writers would be unable to say the kinds of things they like to say. Such words then become easy for speakers to use as well. If an academic speaker uses a word like "categorization" or "development," for example, it may not be to coalesce into a larger intonation unit a nominalization of the verbs "categorize" or "develop," but simply to refer to a well established technical concept. It is possible that the concepts of categorization or development would not have arisen historically without the possibilities provided by writing. However, the current use of such nouns by speakers no longer shows a productive act of nominalizing, but only the use of a technical term. It may be for this reason that the occurrence of nominalizations is relatively high in spoken language, and especially high in lectures, where again we see a convergence with the frequency in letters.

A third device of a similar sort is the preposed or attributive adjective, two occurrences of which can be seen in:

(16) It allows an *unambiguous* manipulation within the *semantic* level.

Although they are less common, attributive nouns play a similar role. An example is the noun "target" in:

(17) Subjects searched for instances of *target* categories.

The distribution of attributive adjectives and nouns together is shown in Table 10. There is little difference between these figures and those given in Table 9 for nominalizations, and similar explanations may apply.

Table 10: Attributive Adjectives and Nouns

Conversations:	23
Lectures:	56
Letters:	55
Academic Papers:	77

To summarize: (1) prepositional phrases, nominalizations, and attributives are commonly used devices in all kinds of language; (2) they are especially frequent in academic writing; and (3) they occur with equal frequency in lectures and in letters. Since they are so frequent in language of every sort, they must be easy-to-use grammatical devices which place no undue cognitive strain on the language producer. Since they are especially frequent in academic writing, they must lend themselves especially well to the kind of intonation unit expansion which takes place in that style. And since they occur with equal frequency in lectures and letters, the opposing goals of lecturers to imitate academic prose and of letter writers to imitate conversation must converge on a single result.

These three are not the only devices whose effect is to increase the length of intonation units. There are a number of others, less frequent, which show a distribution more like that in Table 6, and which thus help to contribute to the overall greater length of intonation units in letters as opposed to lectures. One is the use of "and" to conjoin two elements into a compound phrase, as in the following:

(18) ... slang for patients reflects responses to their *suffering and illness* (paper)

(19) (she) ... tried to help the children *focus and structure* their discourse. (paper)

Although it would not seem particularly difficult to accomplish conjoining of this kind, the fact is that speakers do not do it nearly as often as writers, and that academic writers do it three times as often as conversationalists, as shown in Table 11.

Table 11. Conjoining

Conversations:	8
Lectures:	12
Letters:	18
Academic Papers:	24

A similar distribution is found in the use of present and past participles.[2] Both types may occur either preposed or postposed. Thus, in (20) we find a preposed present participle, and in (21) a preposed past participle:

(20) It was a *recurring* classroom activity.

(21) *Bowed* oscillations can begin in two ways.

In (22) we find a postposed present participle, and in (23) a postposed past participle:

(22) Cult activity *originating* in the hamlets is atypical.

(23) Where they still embody the properties *denoted* by the words which refer to them,

The pooled distribution of these four kinds of participles is shown in Table 12. Here we see a distribution similar to that in Table 11, except that language other than academic writing makes considerably less use of participles, and lecturers fail to use them much more than conversationalists.

Table 12. Participles

Conversations:	5
Lectures:	6
Letters:	11
Academic Papers:	24

To summarize this entire section, a principle way in which written language differs from spoken is in the greater length of written intonation units. Intonation units are expanded with a variety of devices, among the more common of which are prepositional phrases, nominalizations, and attributive adjectives and nouns. Less frequently used devices are the conjoining of words with "and," and participles of several kinds and uses. All of these intonation unit expansion devices are much more frequent in academic writing than in conversations, with lectures and letters being of intermediate frequency. None of these devices is cognitively difficult, but com-bining them in quantity evidently requires the extra time and care available to a writer.

SENTENCE CONSTRUCTION

Conversational spoken language consists in large part of intonation units joined together in a chain, very often with the coordinating conjunction "and" linking them together. In other words, there is a strong tendency for casual speakers to produce simple sequences of coordinated clauses, avoiding the more elaborate interclausal relations found in writing. Elaborate syntax evidently requires more processing effort than speakers can ordinarily devote to it. The following is a typical example of clause coordination in conversational language:

(24) ... And there was two women,
hiking up ahead of us,
... and you sort of got to a rise,
and then the lake was kind of right there
where we were gonna ... camp.
... And the two of them,
.. got to the rise,
and the next minute,
... they just ... fell over,
totally.

A simple way to measure the degree of coordination in a language sample is by counting the coordinating conjunctions (usually "and," less often "but" or "so") which are located at the beginnings of intonation units. We can express this measure as the percentage of intonation units which begin in this way. The percentages for our data are shown in Table 13. Conversational spoken language thus exhibits by far the largest amount of intonation unit coordination, but even lectures show twice as much as letters. Papers show very little in comparison with the other styles. It appears that spoken language of any kind tends to chain clauses together significantly more often than any kind of written language, and that academic writers avoid this practice considerably more than others.

Table 13. Coordination

Conversations:	34
Lectures:	21
Letters:	12
Academic Papers:	4

As a speaker proceeds with the chaining of intonation units, every so often he will end one of these units with the kind of falling pitch associated with the end of a sentence. The function of sentences in spoken language is problematic, but speakers appear to produce a sentence-final intonation when they judge that they have come to the end of some coherent content sequence (Chafe, 1980). What produces this coherence may vary from one instance to the next. It is difficult to predict when a speaker will decide that a chain of intonation units worthy of recognition as a complete sentence has been produced. Speakers are sloppy in this respect, often producing a sentence-final intonation before they mean to, or neglecting to produce one when they should. Premature sentence closures are often followed by afterthoughts; delayed ones give the effect of run-on sentences. The following example is typical:

(25) ... And she said ... um--... she said ... um--
... I got my degree at Harvard,
.. and one of the things we had to do,
... was-- to read.
... Constantly she said.
I read,
... eight hours a day.
.. For four years.
... And I read everything.
.. She said.

And t .. to graduate from Harvard,
... one of the things we had to do was type.
... A paragraph.
.. We had to ... we had to date it,
.. we had to t .. tell who wrote it,
... and she said not only .. do I know
... she said I'm I'm trained to detect ... minute
.. switches in style.

Sentences in written language are better planned than this, giving evidence of the time and effort that went into their construction. We have been led to the conclusion that intonation units are the natural unit of speech, their content and structure dependent on the capacity of short-term memory, whereas sentences have become the major unit of writing. Writers have time to give thought to how much they want to put into a sentence--how best to sculpt it into a complex, integrated whole, as in the following:

(26) A novelist's insistence on the referential function of her work's language is, at least in part, always necessarily divided against itself (as the deconstructionist critics show) because that insistence stems, paradoxically, from an awareness of the novel's fictionality.

One piece of evidence for this difference between spoken and written sentences is the fact that in conversational language the largest number of sentences consist of one word, a slightly smaller number of two words, and so on, with a gradual tailing off to a small number of sentences with a length of 100 words or more. Although the mean length of spoken sentences is 18 words, this figure has no special significance. There is no reason to think that speakers have any stake in producing sentences approximately 18 words long. Academic writing, in contrast, shows a relatively normal distribution of sentence lengths centered around a mean of 24. We take this as an indication that writers possess an intuitive concept of "normal sentence length" which speakers do not have.

In summary, speakers chain intonation units together, often connecting them with "and," stopping occasionally to insert a sentence-final falling pitch which often turns out to have been a mistake. Writers connect intonation units in more complex ways, sculpting them into planned sentences which, in spite of deliberate variability in length, tend toward a mean length of about 24 words. Intonation units are the natural units of speaking, sentences the natural unit of writing, presumably because writers have the leisure to perfect the complex and coherent sentence structures which speakers are moving too fast to produce.

INVOLVEMENT AND DETACHMENT

So far we have looked at properties of speaking and writing which can be attributed to differences in the two activities, and specifically to the rapidity and evanescence of speaking as opposed to the deliberateness and editability of writing. Another, equally important, difference has to do with the relation between the producer of the language and the audience. In most spoken language an audience is not only physically present,

but also has the ability to respond with language of its own. For writers the audience is usually unseen, and often unknown. Spoken language contains indications of the speaker's involvement with the audience, as well as of the speaker's involvement with himself, and furthermore of his involvement with the concrete reality of what is being talked about. Much written language, on the other hand, lacks involvement of any of these three kinds, and is apt to show indications of the writer's detachment from the audience, from himself, and from concrete reality. However, linguistic features of involvement and detachment are not cognitively, but contextually determined. They can, therefore, be overridden when the context is appropriate. They do not necessarily divide spoken language from written, and we will see that it is possible for some written language to be more spokenlike than any spoken language in these respects.

One obvious measure of involvement with an audience is the occurrence of language which responds to something just said by another person. For example:

(27) (Would you do that?)
... Yeah .. I think I would.

(28) (Beatrice said that?)
... That's what I thought she said.

Another such measure is the occurrence of the phrase "you know," with which a speaker attempts to reassure himself that he is getting through to the listener without explicitly requesting confirmation of that fact (cf. Ostman 1982):

(29) So you can't ... *you know* .. fudge in some styrofoam
.. or something like that.

(30) Only this all takes .. *you know* .. ten minutes to
... to compose,

The distribution of both interactional responses and "you know" in our four styles of language is shown in Table 14.

Table 14. Involvement with the Audience

	responses	"you know"
Conversations:	9	11
Lectures:	2	2
Letters:	0	0
Academic Papers:	0	0

For obvious reasons such devices hardly occur at all in either kind of written language, but they are also relatively rare in lectures, whose language, though it may occasionally be interactional, is more often monologic.

A different kind of involvement is explicit concern with oneself, an obvious measure of which is the use of first person pronouns ("I," "me," "my," "mine," "we," "us," "our," and "ours"). The following examples are both from letters:

(31) *I* could be telling you the same things in person.

(32) As you can see, *my* difficulty with writing extends not only to its phrasing but also to its timing.

Table 15 shows the occurrence of first person pronouns in the four kinds of language. They occur very little in academic writing, a fact which must be due in part to their deliberate discouragement in that medium, but in part also to the fact that academic writing usually has a subject matter which excludes much talk about oneself. The most interesting finding in Table 15, however, is that personal letters show more ego involvement than any of the other kinds of language studied here. While first person

Table 15. First Person Pronouns

Conversations:	48
Lectures:	21
Letters:	57
Academic Papers:	4

reference is high in conversations, and even present to a significant degree in lectures, it is by far the highest in letters. If people have a natural inclination to talk about themselves, it is apparently in letters that they have the best oppor-tunity to do so, being freed of whatever inhibitions might be imposed by the immediate presence of an interlocutor. The use of first person pronouns is thus not necessarily a feature which differentiates spoken from written language, but rather a feature which the absence of a direct audience may even foster when the circumstances are right. At the same time, as we can see from the figure for academic papers, writing can create a context in which maximum suppression of one's own identity is possible.

There are numerous linguistic indications of involvement with concrete reality. One that occurs frequently and with an interesting distribution is the use of temporal and spatial adverbs and adverbial phrases. Elements of this kind locate the people and events being talked about in specific time and space. We interpret their use to be one important manifestation of a speaker or writer's involvement with the reality of his subject matter, as opposed to a stance of abstract detachment. Note the occurrence of temporal and locative adverbial elements in the following examples, again from letters:

(33) I applied for this postdoc *last March,*

(34) It looks like life is going to really be nice *in Akron.*

The distribution of such elements in our data is shown in Table 16. They present a picture much like that of Table 15, where the least frequent occurrence is in academic writing and the most frequent in letters. Thus, like involvement with oneself, involvement with specific time and space is not a feature which necessarily differentiates spoken from written language, but rather something which is suppressed in one kind of writing while it is exaggerated in another. Although such involvement may be characteristic of spoken language, and something written language can and often does suppress, there is no processing constraint which prevents it from surfacing and even from predominating in writing when the context is appropriate.

Table 16. Adverbial Expressions

	Locative Adverbials	Temporal Adverbials
Conversations:	14	16
Lectures:	11	10
Letters:	19	22
Academic Papers:	6	8

We use the term detachment as the opposite of involvement. Instead of showing a concern for the concrete aspects of language interaction and for concrete reality, features of detachment show an interest in ideas that are not tied to specific people, events, times, or places, but which are abstract and timeless. We have looked at several features which we take to be manifestations of such detachment. They predominate in academic writing, where they represent the opposite side of the coin from the involvement features just discussed.

Involved language tends to favor clauses whose subjects refer to specific concrete persons, usually persons who perform some concrete action or are in some concrete state. Detached language uses clauses whose subjects refer to abstractions. For example (from academic papers):

(35) *This suggestion* finds some support in studies of children's "egocentric speech"

(36) *Educational settings* are rich in such situations.

The two subjects, "this suggestion" and "educational settings," illustrate the kinds of abstractions typical of such language. The distribution of abstract subjects in our data is shown in Table 17. As might be expected, their occurrence is lowest in conversations, but it is almost as low in letters, whose penchant for concreteness we have already noticed. While lectures show a higher degree of this kind of abstractness, they do not approach the frequency to be found in academic writing.

Table 17. Abstract Subjects

Conversations:	21
Lectures:	38
Letters:	25
Academic Papers:	48

A similar feature, though less common throughout our samples, is the use of passive constructions. The use of a passive allows a writer to avoid mentioning any concrete doer, and in that way to treat an event in a more abstract fashion. Even when an agent is mentioned, a passive construction shows less concern for concrete people doing concrete things, in that it gives subject status to things (usually abstractions) which have something done to them. Examples from academic papers are:

(37) The resonance complex *has been studied* through experiments with an electronic violin.

(38) This tendency *has been remarked upon* by other researchers.

By far the largest number of passives occurred in our samples of academic writing, as shown in Table 18. There were very few in conversations, and only a modest number in either lectures or letters. It would seem that the passive construction provides a device that academicians find especially suited to their purpose, while no one else is especially fond of it.

Table 18. Passives

Conversations:	3
Lectures:	9
Letters:	7
Academic Papers:	22

Still another indication of detachment is the use of words which express the probability of some generic statement being true. We have found that academic writers are particularly fond of expressions which indicate that things happen, in general, a certain proportion of the time, but not necessarily always--words like "normally," "usually," "primarily," "principally," and "virtually." Such words are academic hedges, by which the writer escapes blame for instances which fail to correspond to his generalization. Some examples are:

(39) Correction is *usually* thought of as being only a matter of stopping and saying part of your sentence over differently.

(40) The contour used *primarily* by the white children was a gradually rising contour.

Such expressions occur noticeably more often in academic writing than elsewhere, as shown in Table 19.

Table 19. Indications of Probability

Conversations:	4
Lectures:	7
Letters:	6
Academic Papers:	10

In summary, different kinds of language show different degrees of involvement and detachment. Involvement with one's audience is obviously an important aspect of conversations. It is a minimal feature of lectures, and is lacking in any kind of writing. Involvement with oneself, though conspicuous in conversations also, is present

to an even greater degree in letters, a fact which suggests that letters provide the best forum of all for egocentrism. Involvement with concrete reality, as measured in terms of temporal and locative adverbs, shows a similar pattern. Letter writers have the time for such expression, and lack the inhibitions against it shown by academic writers. Academic writing shows various manifestations of detachment from concrete reality, including maximum use of abstract subjects, of passives, and of probabilistic generalizations.

CONCLUSION

Language users, whether they are speakers or writers, have certain variable resources available to them. Every speaker or writer possesses a repertoire of devices which are combined in varying mixtures depending on the context, the purpose, and the subject matter of language use. In other words, language adapts to its varying environments. It may be useful for us to summarize the varying characteristics of spoken and written language as they have appeared in this survey.

Conversationalists, first of all, employ a relatively limited vocabulary, and they are inclined to hedge their lexical choices and to be referentially inexplicit. They make considerable use of colloquial words and phrases. They create relatively brief intonation units, which they chain together, stopping every so often to make a sentence boundary which is not always well justified in terms of topical coherence. They interact with their audiences, show ego involvement, and talk frequently about specific times and places. These properties are appropriate to language that is produced rapidly in an environment where the immediate presence of the audience plays an important role.

Academic lecturers speaking in relatively informal contexts employ an equally limited vocabulary, also use hedges, and are also referentially inexplicit. Their use of literary vocabulary is somewhat greater than that of conversationalists. They are, however, only slightly more able to extend the length of intonation units, and they are no more able than conversationalists to cope with syntactically elaborate sentences. They interact slightly with their audiences, but not nearly as much as conversationalists. They make some use of first person and concrete spatio-temporal references, though not as much as either conversationalists or letter writers. Conversely, they show some degree of detachment, but not as much as academic writers. This is a mixed kind of language, still controlled by the constraints of rapid production, but striving after some of the elegance and detachment of formal writing.

Letter writers use a more varied vocabulary, are sometimes inexplicit, but use hedges very rarely. They use a moderate number of colloquial words and contractions, but at the same time a greater number of literary items. Their intonation units are intermediate in length between conversationalists and academic writers, and their sentences tend to be well formed. They show a greater degree of involvement with themselves and with concrete reality than any other language users within the range that we have studied. Their degree of detachment is not much greater than that of conversationalists. This is a kind of language which takes some advantage of what the deliberateness of writing allows, but which also maintains, and in some respects even surpasses, the casualness and involvement of speech.

Academic writers represent for us the extremes of what writing permits. Their vocabulary is maximally varied, and they avoid both hedges and inexplicit references. Their writing is maximally literary, with almost no colloquial items or contractions.

Their intonation units are maximally long, with frequent use of all the devices by which intonation units can be expanded. Their sentences show a tendency toward an ideal length, and are maximally coherent. They show little involvement with themselves, or with concrete reality, and on the contrary make the greatest use of devices signaling detachment. This kind of language represents a maximum adaptation to the deliberateness and detachment of the writing environment.

This paper has dealt only with the most frequent of the linguistic features we have examined. We will have more to say in the future about various other features whose occurrence is less pervasive. Beyond that, we need to emphasize that the four types of language we have discussed here are only four out of many. We expect in the future to deal not only with literary and journalistic language, for example, but also with several styles of spoken and written language produced by third and sixth grade children. Our preliminary findings have suggested that as children learn to write, they first learn to slough off the most positive features of spoken language, those which we have presented here under the heading of involvement, so that there is a stage at which their writing can best be characterized as bland. It is only later and more gradually that they begin to gain proficiency in handling the positive features of written language, above all the elaboration of sentences. It will be interesting to see how well this two stage sequence of writing acquisition is supported by additional data.

In the meantime, we hope to have provided a start toward a better understanding of the ingredients of language which both speakers and writers use in differing combinations for different purposes. A general hope is that dissemination of research of this kind will lead teachers to recognize more clearly, not only that writing and speaking are very different activities, but also that the kinds of language which result from these activities are very different too, and that there are good cognitive and social reasons for such differences. We hope to have provided some specific clues to what the linguistic differences are, as well as to the factors which are responsible for them.

Footnotes

[1]We also collected roughly parallel samples from 20 third graders and 20 sixth graders, but they will not be discussed here. This project was sponsored by Grant G-80-0125 from the National Institute of Education. We are grateful to Pamela Downing for her valued assistance and advice.

[2]We exclude here the use of participles in progressive and perfect constructions like "he is writing" and "he has written," as well as in passive constructions like "it has been written".

References

Bloomfield, L. *Language*. New York: Holt, 1933

Chafe, W. The deployment of consciousness in production of a narrative. In W. Chafe (Ed.), *The pear stories: Cognitive cultural, and linguistic aspects of narrative production*. Norwood, J.J.: Ablex, 1980.

Chafe, W. Differences between colloquial and ritual Seneca, or how oral literature is literary. In A. Schlichter, W. Chafe, and L. Hinton (Eds.), *Reports from the Survey of California and Other Indian Languages, No. 1*. Department of Linguistics, University of California, Berkeley, 1981.

Chafe, W. Integration and involvement in speaking, writing, and oral literature. In Tannen, 1982(a).

Chafe, W. Linguistic differences produced by differences between speaking and writing. In D. R. Olson, A. Hildyard, and N. Torrance (Eds.), *Literacy, language, and learning: The nature and consequences of reading and writing*. Cambridge: Cambridge University Press, 1985.

Chafe, W. Evidentiality in English Conversation and Academic Writing. In W. Chafe and J. Nichols (Eds.), *Evidentiality: The linguistic coding of epistemology*. Norwood, N.J.: Ablex, 1986.

Goody, J., and I. Watt. The consequences of literacy. In J. Goody (Ed.), *Literacy in traditional societies*. Cambridge: Cambridge University Press, 1968.

Kroll, B. M., and R.J. Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts*. Urbana: National Council of Teachers of English, 1981.

Lakoff, G. Hedges: A study in meaning criteria and the logic of fuzzy concepts. In D. Hockney et al. (Eds.), *Contemporary research in philosophical logic and linguistic semantics*. Dordrecht: D. Reidel, 1975.

Olson, D. R. From utterance to text: the bias of language in speech and writing. *Harvard Educational Review* 47:257-281, 1977.

Ong, W. *Orality and literacy: The technologizing of the word*. London and New York: Methuen, 1982.

Ostman, J.-O. *You know: A discourse-functional approach*. Amsterdam: John Benjamins, 1981.

Pawley, A., and F. Syder. Sentence formulation in spontaneous speech: the one-clause-at-a-time hypothesis. Paper read at the First Annual Congress of the Linguistic Society of New Zealand, Auckland, August, 1976.

Scribner, S. and M. Cole. *The psychology of literacy*. Cambridge: Harvard University Press, 1981.

Tannen, D. *Spoken and written language: Exploring orality and literacy*. Norwood, N.J.: Ablex, 1982(a).

Tannen, D. Oral and literate strategies in spoken and written language. *Language* 58:1-21, 1982(b).

Vachek, J. *Written language*. The Hague: Mouton, 1973.